To appear in *Molecular Simulation*
Vol. 00, No. 00, Month 2009, 1–17
10.1080/0892702YYxxxxxxxx

Mechanism for asymmetric bias in demonstrations of the NPI and fluctuation theorem

Charlotte F. Petersen,* Denis J. Evans and Stephen R. Williams*

Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

(*Received 00 Month 20XX; final version received 00 Month 20XX*)

We consider two different methods of calculating the relevant average for the Nonequilibrium Partition Identity (NPI), *i.e.* $\langle \exp\left[-\bar{\Omega}_t t\right] \rangle$, which result in two different values. At best only one of these will accurately correspond to what is observed. In order to better understand the two outcomes we carry out a detailed error analysis. This analysis is difficult due to the importance of extremely rare events in forming the average, resulting in the necessity to go beyond linear approximations for the error estimates. We begin by analysing the error in the fluctuation relation, and build upon this to estimate the errors in the NPI average. At short durations the full ensemble average always gives the observed average (i.e. the NPI holds). However, at very long durations, given a fixed amount of sampling, the observed average is predicted by treating the probability distribution as a Dirac-delta function. At intermediate times, neither corresponds to the observed average. This has profound implications for nonequilibrium work relations, as first introduced by Jarzynski.

Keywords: fluctuation theorem; nonequilibrium; molecular dynamics; asymmetric data; uncertainties; nonequilibrium partition identity; integral fluctuation theorems

1. Introduction

Over the past couple of decades we have gone from knowing very few exact results in nonequilibrium statistical mechanics, to the current status where we know a number of exact relations that are valid for systems of arbitrary size, arbitrarily near or far from equilibrium [1–9]. These include fluctuation relations, summing rules (*i.e.* exact results about the values of the ensemble averages of quantities involving nonequilibrium path integrals), the dissipation theorem of nonlinear response theory [7] and a number of equilibrium relaxation theorems, e.g. [9, 10]. Obviously these results, which place new constraints on what can occur far from equilibrium, may require us to revise our view point in certain cases. An relevant example, is the behaviour of the so called dissipative flux when a transient state, driven by a constant field, is observed for a long enough that the system can be accurately modelled as a steady state. Consideration along these lines leads us directly to an interesting and challenging problem about the behaviour of the summing rule given a finite amount of sampling. First however, we must consider the behaviour of the summing rule under a full ensemble average.

Given an exact fluctuation relation there will be a summing rule that can be derived from it. For the Crooks Fluctuation Theorem (CFT)[4, 11] there is the well known Jarzynski Equality summing rule (JE)[12, 13] and for the Evans-Searles Transient Fluctuation Theorem (ESFT)[2] there is the Nonequilibrium Partition Identity summing rule (NPI)[14]. If we express the former case (CFT &

*Corresponding authors. Email: charlotte.petersen@anu.edu.au, swilliams@rsc.anu.edu.au

JE) in terms of the purely irreversible work, it then becomes mathematically the same as the NPI. This summing rule places a constraint on the statistical fluctuations in the steady state. We need to reconcile this with the established physical behaviour that the material must follow.

The transient Evans-Seales fluctuation theorem (ESFT) is one of the important exact results to be obtained in non-equilibrium statistical mechanics. It gives the ratio of observing opposite values for the time averaged dissipation function:

$$\frac{Pr(\bar{\Omega}_t = A)}{Pr(\bar{\Omega}_t = -A)} = \exp[At], \tag{1}$$

where t is the duration for which the time average is evaluated over, $\bar{\Omega}_t$ is given by a path integral in the form of the time averaged dissipation function

$$\bar{\Omega}_t t \equiv \Omega_t \equiv \int_0^t ds\, \Omega(s), \tag{2}$$

where Ω_t is defined in terms of the initial probability distribution, $f(\mathbf{\Gamma}, 0)$, (which is assumed to be an even function of the momenta) as:

$$\Omega_t \equiv \ln\left[\frac{f(\mathbf{\Gamma}(0), 0)}{f(\mathbf{\Gamma}(t), 0)}\right] - \int_0^t ds\, \Lambda(\mathbf{\Gamma}(s)). \tag{3}$$

In this equation $\mathbf{\Gamma}(\mathbf{p}, \mathbf{q})$ is the phase space vector representing all the momenta and all the positions of all the particles comprising the system, $\mathbf{\Gamma}(t)$ is the phase space vector at time t as evolved to by the equations of motion starting from $\mathbf{\Gamma}(0)$ at time 0, $f(\mathbf{\Gamma}, t)$ is the phase space density (the ensemble probability function) at time t and $\Lambda(\mathbf{\Gamma}) = \partial\dot{\mathbf{\Gamma}}/\partial\mathbf{\Gamma}$ is the phase space expansion factor given by the divergence of the equations of motion.

We can now obtain the NPI (the summing rule resulting from the ESFT) by integration of the ESFT,

$$\int_{-\infty}^{\infty} dA\, Pr(\bar{\Omega}_t = A) \exp[-At] = \int_{-\infty}^{\infty} dA\, Pr(\bar{\Omega}_t = -A). \tag{4}$$

Since the probability distribution $Pr(\bar{\Omega}_t = A)$ is normalised, we have the NPI

$$\left\langle \exp[-\bar{\Omega}_t t] \right\rangle = 1. \tag{5}$$

Let us now consider a specific example to illustrate the point. When we drive a system by a constant, purely dissipative field and when the initial distribution function $f(\mathbf{\Gamma}, 0)$ is the equilibrium distribution function, the dissipation function is very closely related to the dissipative flux, which will be related to a material property of the system. Mathematically this is expressed as [15]

$$\Omega(t) = -\beta J(\mathbf{\Gamma}(t)) V F_e, \tag{6}$$

where $\beta = 1/(k_B T)$, k_B is Boltzmann's constant, T is the equilibrium temperature the system will relax to if it is so allowed, V is the system volume and F_e is the dissipative field. We now consider the case of planar Couette flow where the material property will be the viscosity, which is a transport coefficient. The equations of motion for planar shear result in [16]

$$J(\mathbf{\Gamma}) V F_e = P_{xy}(\mathbf{\Gamma}) V \dot{\gamma} \tag{7}$$

where P_{xy} is the xy element of the pressure tensor, or the negative of the shear stress, and $\dot{\gamma}$ is the strain rate. In the steady state the nonlinear viscosity is given by

$$\eta(\dot{\gamma}) = -\frac{\langle P_{xy} \rangle}{\dot{\gamma}} \tag{8}$$

and in the limit where the shear rate goes to zero this gives us the Newtonian shear viscosity. It is now a simple matter to relate the dissipation function, as it appears in the NPI, to the viscosity, i.e.

$$\eta(\dot{\gamma}) = \frac{k_B T}{V \dot{\gamma}^2} \lim_{t \to \infty} \langle \bar{\Omega}_t \rangle \,. \tag{9}$$

Now we are in a position to see how our expectations regarding the material properties with regard to the viscosity must be interpreted in a way that satisfies the NPI, as promised above. Given large enough values for t and N_p, the number of particles, the standard deviation for the time average $\bar{\Omega}_t$ will scale as $\sqrt{N_p/t}$. So in the limit of infinite time we expect that its standard deviation should approach zero and the distribution function should be accurately represented by a Dirac-delta distribution. The mean value of this distribution is easily found by referring to Eq. (9) and given in terms of the viscosity,

$$\lim_{t \to \infty} \langle \bar{\Omega}_t \rangle = \beta V \dot{\gamma}^2 \eta(\dot{\gamma}). \tag{10}$$

So we may now evaluate the NPI given this knowledge about the Dirac-delta distribution that leads us directly to

$$\langle \exp[-\bar{\Omega}_t t] \rangle = \int_{-\infty}^{\infty} dA \, \delta(A - \beta V \dot{\gamma}^2 \eta(\dot{\gamma})) \exp[-At] = \exp[-\beta V \dot{\gamma}^2 \eta(\dot{\gamma}) t], \tag{11}$$

which clearly isn't equal to the value of unity as required by the NPI. How do we reconcile this with the NPI? The answer is straightforward. The derivation of the NPI relies on the fact that given a finite likelihood of observing the value of $\bar{\Omega}_t = A$ there is also a finite likelihood of observing the value of $\bar{\Omega}_t = -A$ as specified by the ESFT, Eq. (1), and that both of these paired values must be summed. In fact the definition of the dissipation function, Eq. (3), requires that if $f(\Gamma, 0) \neq 0$, $f(M^T \Gamma(t), 0) = f(\Gamma(t), 0) \neq 0$, where M^T is the time reversal mapping: $M^T(q, p) \equiv (q, -p)$. This is known as the ergodic consistency condition. If it is violated the time integrated dissipation function given in Eq. (3) is undefined. Assuming that the distribution function may be represented by a Dirac-delta function violates ergodic consistency, and this leads to a result that is different to that given by the full ensemble average.

We also need to distinguish two types of ergodic consistency. Is the system ergodically consistent in principle or is the data that is accessible to us on any reasonable timescale of observation, ergodically consistent? Although we may in principle be able to observe the antitrajectories, in practice if the probablitity of observing these antitrajectories is negligible over any available timescale, almost any data set we can observe may violate ergodic consistency. Moving from the domain where ergodic consistency in the data sets is satisfied to the regime where it is almost never satisfied changes the way we have to model the experiment. In the former regime the NPI can be expected to be satistfied while in the latter regime it will almost never be satisfied and instead the results will be modelled using equations such as (11). The transition region between the two regimes will be a main subject of investigation in this paper.

Despite the tremendous interest in these types of summing rules over the past 10-15 years [5, 6, 17–30] there has been no analysis of the type to be undertaken here, showing in detail

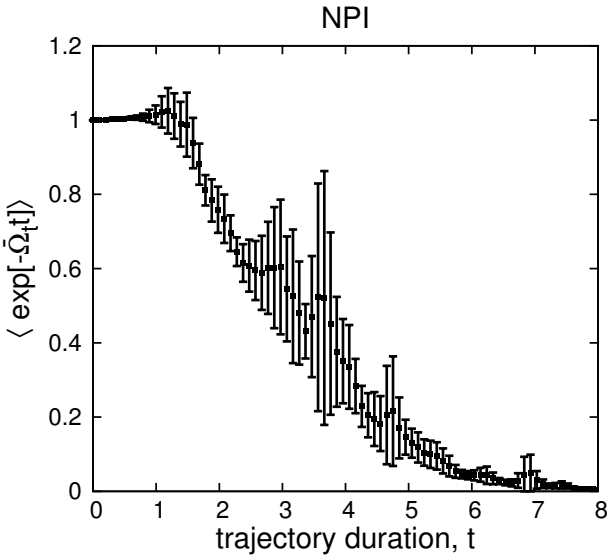

Figure 1. NPI calculated for an ensemble as the duration of the trajectory is increased.

how the errors due to rare events accumulate and even come to dominate the ensemble average. Determining the magnitude of the uncertainty is inherently difficult because the distribution of errors around the mean is not Gaussian and given enough duration they will be dominated by rare events.

As a prelude we show the results for estimates of the NPI obtained from nonequilibrium molecular dynamics simulations, Fig. 1. Here the process of colour conductivity rather than planar shear has been used with a field strength of $F_e = 2$. For now the change of process is of no importance, none the less it will be fully detailed later in the paper. There are a number of important features that can be seen in this graph. Firstly the NPI works quite accurately up to a duration of approximately $t \approx 1$, where it can be seen to sum to its expected value of unity. The graph also has error bars estimated from 10 separate block averages representing 2 standard errors. These error bars appear quite reasonable up to a duration of around $t = 1.5$, but at longer durations the errors estimated using the block averages are hopelessly inadequate.

In addition we can observe a strong serial correlation in this data as the duration t is varied, due to trajectory segments starting from the same initial phase space points, $\mathbf{\Gamma}(0)$ being used for all durations t shown in the figure. As the duration increases the rare events become not just rarer but in compensation they each make a stronger contribution to the full ensemble average value of unity. This combined with the serial correlation results in the NPI being significantly and consistently underestimated at long enough durations, because the rare events due to the backward flow are so rare that they have become completely unobservable. From a practical point of view the feature in the graph of the estimates substantially underestimating the full ensemble average value of unity is repeatable, resulting in the overly small error bars from the standard error analysis. It is also interesting to note that when a data point happens to have a larger value for the average relative to the neighbouring data points, the error bars are consistently larger for that point. This is due to the increased sampling of rare events, which increase the inferred variance used to obtain the standard error displayed in the figure. These are the general features that arise upon applying a limited amount of serially correlated data to the summing rules that result from fluctuation relations.

2. Using the Fluctuation Theorem to Estimate Uncertainties in the Frequency Histogram

The NPI is a consequence of the ESFT, so understanding the statistical fluctuations in the ESFT is a good place to begin before tackling the more difficult case of the NPI. As it turns out, we can build upon the results from the analysis of ESFT in order to analyse the NPI. Demonstrating the ESFT involves comparing the number of trajectories which have a dissipation function value of $\bar{\Omega}_t = A \pm \delta A$ to the number with the less likely value, $\bar{\Omega}_t = -A \pm \delta A$. Practically, we compare these values by comparing the relative heights of the corresponding bins of a frequency histogram. Often, the counts for the bins in the negative wing of the distribution will be dominated by noise. The error in the ratio of frequencies will be most significant when the negative bin has far fewer counts than the positive bin. To model the error in the bin height we need to know the expected height, which can be determined by mapping the positive side of the distribution to the negative side using the ESFT itself. This will result in a distribution that is as noisy as the positive side of the original distribution. The expected value for each histogram bin is now given by

$$\lambda(A) = \begin{cases} \exp[At] n_m(-A) & \text{if } A < 0 \\ n_m(A) & \text{if } A \geq 0, \end{cases} \tag{12}$$

where $n_m(A)$ is the measured frequency of the histogram bin for $\bar{\Omega}_t = A \pm \delta A$. In systems that have an approximately Gaussian distribution of dissipation function values, we can estimate the true height of the bin by replacing the distribution with a Gaussian for the purpose of calculating the uncertainty. The mean of this distribution is set to the mean of the data. The variance of the Gaussian distribution that satisfies the ESFT is twice the mean[31]. This would result in less noise in each expected bin height. For the rest of the analysis we will consider the more general case where a Gaussian approximation cannot be used.

3. Binomial Distribution

The binomial distribution gives the probability distribution for a biased coin toss experiment. That is, for a process with a fixed probability of success, r, performed N times, the probability of obtaining exactly k successes is given by the binomial distribution,

$$p_k(r, N) = \binom{N}{k} r^k (1-r)^{N-k}. \tag{13}$$

The height of each histogram bin in our system, $n(A)$, can be described by a binomial distribution where N is the total number of trajectories that have been simulated and $r = \lambda(A)/N$ is the probability of a trajectory falling within a particular histogram bin.

This distribution function is related to the frequency histogram used to demonstrate the ESFT in the following way. There are a total of N trajectories computed, with the expectation value $\lambda(A)$ for the number of these trajectories found in the histogram bin $A \pm \delta A$. So the binomial distribution gives the probability of observing a value of k trajectories falling in the histogram bin given by $A \pm \delta A$. Note that the binomial distribution Eq. (13) is skewed and asymmetric, particularly for distributions with lower values for the average, λ.

To generate error bars from this distribution, which gives the probability of a given number of events, we need to use the cumulative probability distribution, $CDF_k(r, N)$, where k is the highest number of occurrences that the probability mass function is summed to.

To get error bars of certain confidence we can calculate the value of k required to give specific

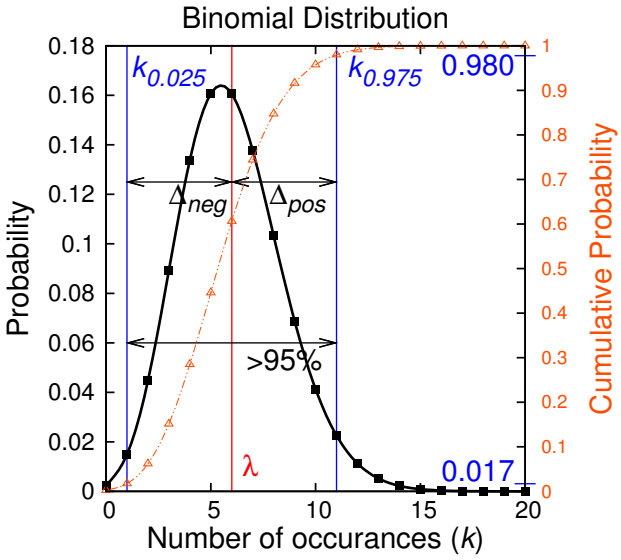

Figure 2. The binomial distribution function with a mean of 6 and $N = 10^8$ (squares), overlaid with its cumulative distribution (triangles). The size of error bars which would cover at least 47.5% of the distribution on each side of the median are shown.

values of the CDF. For example, to include at least 95% of the distribution, the error bars would span from the value of $k = k_{0.025}$, the largest integer value satisfying $CDF_{k_{0.025}}(\lambda) \leq 0.025$ to the value of $k = k_{0.975}$, the smallest integer value satisfying $CDF_{k_{0.975}}(\lambda) \geq 0.975$. For the histogram bin associated with the variable A (e.g. recall Eq. (12)), we can label the positive and negative error bars as $\Delta_{pos}(A) = k_{0.975}(A) - \lambda(A)$ and $\Delta_{neg}(A) = \lambda(A) - k_{0.025}(A)$ respectively. In general we have

$$\Delta_{pos}(A) = k_{0.5+p/2}(A) - \lambda(A) \tag{14a}$$

$$\Delta_{neg}(A) = \lambda(A) - k_{0.5-p/2}(A), \tag{14b}$$

where p is the proportion of the histogram to be included in the error bars, and the value of k is determined using the CDF. These error bars may be asymmetric due the skew of the distribution. An example of how these error bars are calculated can be seen in Figure 2.

The average number of occurrences in the histogram bin will be different for each bin, and so will have a different corresponding binomial distribution. This means that the error bars will vary with each bin, in both size and symmetry. The binomial distribution is more skewed for smaller values of λ, resulting in relatively less symmetric error bars for less populated bins. We will only consider histogram bins where the error bars do not overlap with zero, to avoid dividing by zero. In these cases, the calculated error bars are roughly symmetric.

4. Numerical System

To illustrate the method we will use data from a model computational system. To generate a nonequilibrium system we will apply a weak colour field to all the particles in the system[2], as

mentioned previously in regard to Fig. (1). The equations of motion used are

$$\dot{\mathbf{q}}_i = \frac{\mathbf{p}_i}{m}, \tag{15a}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i + c_i\mathbf{F_e} - \alpha\mathbf{p}_i, \tag{15b}$$

where $\mathbf{q}_i$ is the position of particle i, $\mathbf{p}_i$ is the momenta, m is the particle mass, $\mathbf{F}_i$ is the interparticle force determined by the WCA potential[32], $c_i = (-1)^i$ is the colour label and $\mathbf{F_e}$= $(F_{ex}, 0)$ is the colour field in the x direction with strength F_{ex}. An isokinetic thermostat[16] is used, $\alpha = (\sum_i^{N_p} \mathbf{F}_i \cdot \mathbf{p}_i + \sum_i^{N_p} c_i\mathbf{F_e} \cdot \mathbf{p}_i)/(\sum_i^{N_p} \mathbf{p}_i \cdot \mathbf{p}_i)$. The system begins in an equilibrium canonical distribution. This was generated with an equilibrium molecular dynamics simulation, with starting points for the transient simulations taken every 2,000 time steps. This ensures the simulations are statistically independent. Our simulation is carried out in reduced units. We use a two dimensional system of 8 particles (N_p=8) with a number density of $\rho = N_p\sigma^2/V = 0.6$. The trajectory duration is labelled t, with a time step length of 0.001. $N = 9 \times 10^7$ trials of the simulation were completed. The histogram bin width was initially 0.003, and only bins which contained data are included. We will initially use a field strength of 0.17 and a trajectory duration of 12. The form of the dissipation function for this system is well known[15], and given by $\Omega(t) = \beta F_{ex} \sum_{i=1}^{N_p} c_i\dot{x}_i(t)$.

5. Propagating Uncertainties to Demonstrate the Fluctuation Theorem

To demonstrate the fluctuation theorem we will take the approach of calculating the gradient of the LHS against the RHS, for each pair of histogram bins. That is,

$$g = \frac{1}{At} \ln\left[\frac{n_m(A)}{n_m(-A)}\right]. \tag{16}$$

For the fluctuation theorem to be satisfied, we would expect g to have a value of unity for each histogram bin. Using our example computational system g was calculated for each value of A, shown in Figure 3. To calculate the total slope of the LHS of the fluctuation theorem against the RHS, we can construct a weighed average of the gradient calculated for each pair of histogram bins,

$$G = \frac{\sum_i^h g_i/\left\langle \Delta g_i^2 \right\rangle}{\sum_i^h 1/\left\langle \Delta g_i^2 \right\rangle}. \tag{17}$$

Each point is weighted by the inverse variance, calculated from the variance of the binomial distribution for each histogram bin using the standard error propagation method,

$$\begin{aligned} \left\langle \Delta g_i^2 \right\rangle &= \left\langle \frac{\partial g}{\partial n(A)} \right\rangle^2 \left\langle \Delta n(A)^2 \right\rangle + \left\langle \frac{\partial g}{\partial n(-A)} \right\rangle^2 \left\langle \Delta n(-A)^2 \right\rangle \\ &= \frac{\left\langle \Delta n(A)^2 \right\rangle}{(A\left\langle n(A) \right\rangle)^2} + \frac{\left\langle \Delta n(-A)^2 \right\rangle}{(A\left\langle n(-A) \right\rangle)^2} \end{aligned} \tag{18}$$

The variance of the binomial distribution is given by $\left\langle \Delta n(A)^2 \right\rangle = \lambda(A)(1 - \lambda(A)/N)$, giving us

$$\left\langle \Delta g_i^2 \right\rangle = \frac{1}{A^2}\left[\frac{1}{\lambda(A)} + \frac{1}{\lambda(-A)} - \frac{2}{N}\right]. \tag{19}$$

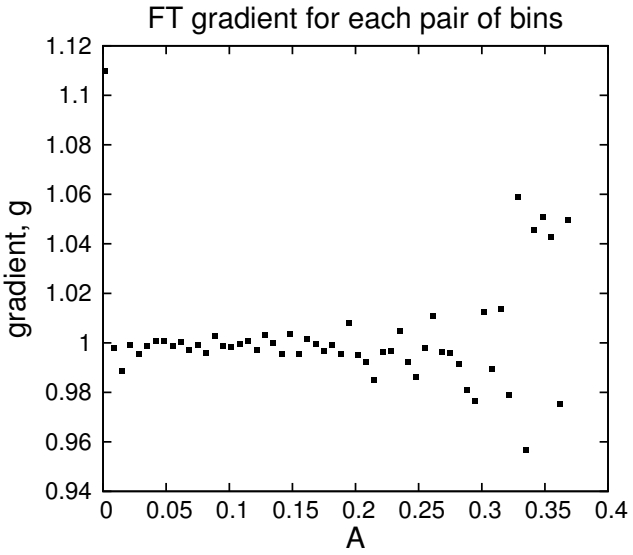

Figure 3. Gradient calculated form each pair of bins, given by Eq. (16). The histogram of dissipation function values was taken from the example computational system.

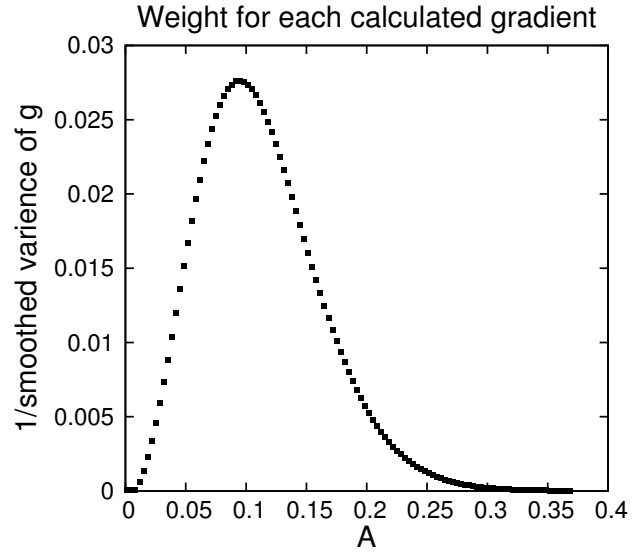

Figure 4. Inverse of the smoothed variance of each point, calculated from the binomial distribution describing each histogram bin's height in Eq. (19).

To weight each point we will smooth the calculated variance, so that each variance is replaced by the average of itself with 2 data points on each side. The inverse of this calculated variance is shown for each pair of histogram bins in Figure 4. Points close to the origin receive a low weight because small fluctuations in their value have a large effect on the slope. Points in the wings of the distribution have a low weight because of the large variance in the histogram bin height relative to their average height. The weighted average of the slope is plotted in Figure 5 as we increase the number of histogram bins included in the average. The average quickly converges to its expected value of unity as more data is added. While the calculated slopes for histogram bins in the wings of the distribution are significantly noisier, this is not reflected in the final average because these

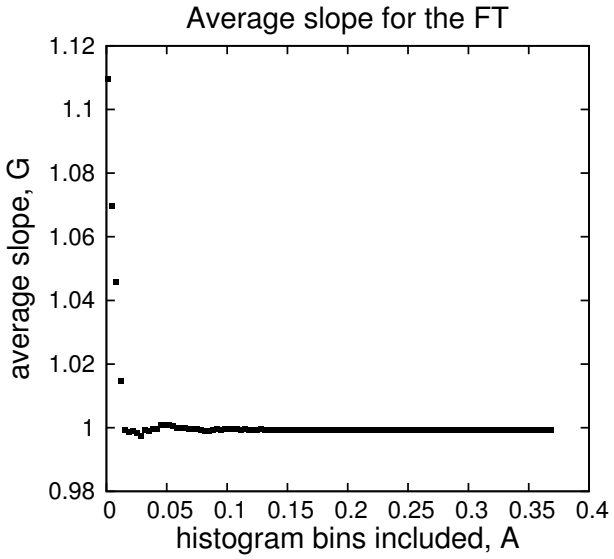

Figure 5. Slope of the ES-FT calculated as a weighted average of slopes calculated from each pair of histogram bins, given by Eq. (17). The average was calculated using all pairs of histogram bins up to the value of A given on the x-axis.

points have low weights.

Propagating the variance from the binomial distribution into the variance for g does not take into account the asymmetry in the initial distribution for each bin height, or any asymmetry due to the small size of the histogram bin height relative to the expected range in this height in the bins near the wings of the distribution. To account for this, we can calculate the largest and smallest possible values for g given that the height of each histogram bin is determined by a specific proportion of the binomial distribution in Eq. (14).

$$\Delta_{g+}(A) = \frac{1}{A} \ln \left[\frac{n(A) + \Delta_{pos}(A)}{n(-A) - \Delta_{neg}(-A)} \right] - 1, \tag{20a}$$

$$\Delta_{g-}(A) = 1 - \frac{1}{A} \ln \left[\frac{n(A) - \Delta_{neg}(A)}{n(-A) + \Delta_{pos}(-A)} \right]. \tag{20b}$$

In doing this we have not assumed that the errors are small enough to be accurately approximated as linear, as was done in Eq. (18). This range of possible values for g is plotted in Figure 6 for each pair of histogram bins in our example system. We can see that for large values of A the range is asymmetric. That is, the slope can land further from the expected value of unity on the positive side than on the negative side. We can propagate these asymmetric ranges to the weighed average of the slope, G, using the standard error propagating method, but applying it separately to the positive and negative range.

$$\Delta_{G+} = \sqrt{\sum_i^h \left(\frac{\partial G}{\partial g_i} \right)^2 \Delta_{g_{i+}}^2} = \frac{\sqrt{\sum_i^h \Delta_{gi+}^2 / \left\langle \Delta g_i^2 \right\rangle^2}}{\sum_i^h 1 / \left\langle \Delta g_i^2 \right\rangle}, \tag{21a}$$

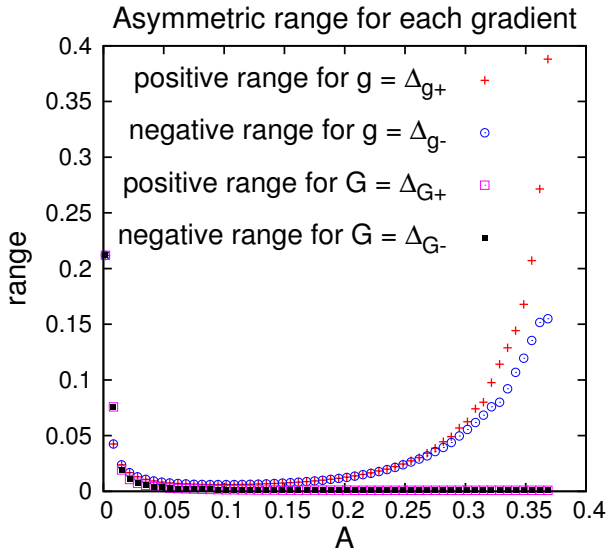

Figure 6. Range of the gradient, as given by the qualities specified by Eqs. (20 & 21), for each pair of bins, calculated in Eq. (20). Also shown is the range for the weighed average of the gradient, calculated in Eq. (21). This range was calculated using all pairs of histogram bins up to the value of A given on the x-axis. Note that these positive and negative ranges appear to overlap.

$$\Delta_{G-} = \sqrt{\sum_i^h \left(\frac{\partial G}{\partial g_i}\right)^2 \Delta_{g_{i-}}^2} = \frac{\sqrt{\sum_i^h \Delta_{gi-}^2 / \left\langle \Delta g_i^2 \right\rangle^2}}{\sum_i^h 1/\left\langle \Delta g_i^2 \right\rangle}. \tag{21b}$$

This range of values is included in Figure 6. The range decreases as the accuracy of G increases with increasing A. While the limits for g were asymmetric for the wings of the distribution, this asymmetry is not observable in the limits for G because of the low weight of these points. The errors given by Eqs. (21) again are assumed linear, but here the errors are (except for small values of A) very small, so this approximation will be accurate.

6. NPI

As a first step towards extending this analysis to the NPI we will calculate the value of the NPI for each pair of histogram bins. We can show using the ESFT that this is expected to have the same value as the full NPI,

$$\left\langle e^{-\bar{\Omega}_t t} \right\rangle_A = \frac{\exp[-At]n(A) + \exp[At]n(-A)}{n(A) + n(-A)} \tag{22}$$

$$= \frac{\exp[-At]Pr(A)N + \exp[At]Pr(-A)N}{n(A) + n(-A)} \tag{23}$$

$$= \frac{N\left(Pr(-A) + Pr(A)\right)}{n(A) + n(-A)} = 1. \tag{24}$$

We calculated this in the same system as before, but with a field strength of $F_{ex} = 0.29$, trajectory duration of $t = 18$, $N = 9\times10^7$ trials and histogram bin width of 0.004. This is plotted in Figure 7 as crosses. The error bars are constructed by splitting the data into 9 sets and calculating the standard error. For histogram bins close to the centre (small values of A), the NPI has a value within the

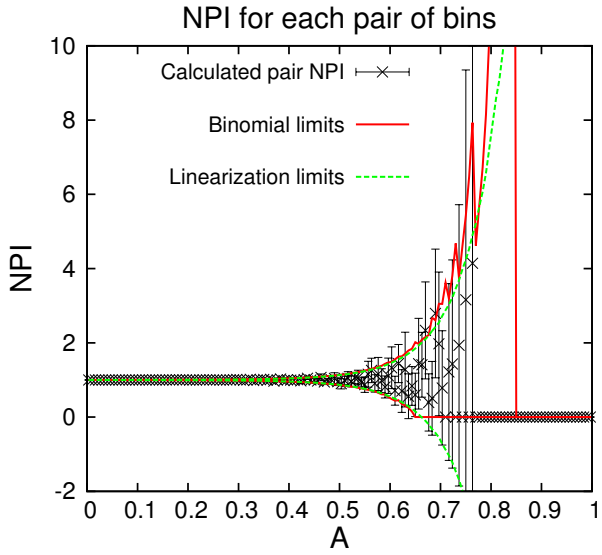

Figure 7. NPI calculated for each pair of histogram bins determined by A, shown as black crosses. The error bars are 2 times the standard error. The green dashed lines show the expected range of the pair NPI calculated from linearising the equation and using the variance of the binomial distribution describing the height of each bin. The limits calculated from the largest and smallest expected heights in each of the histogram bins are shown as the solid red line.

error bars of unity, as expected. As the value of A increases the data becomes significantly noisier, though the error bars still overlap with the value of unity. However, in the wings of the distribution the NPI for a number of histogram bins drops to a value close to zero, with very small error bars. This is very clear for values of $A > 0.75$. In these cases, the histogram bin corresponding to $-A$ did not record any trajectories, and we can see that the standard error analysis method has failed. It is when these histogram bins make a significant contribution to the total NPI that its value appears to be significantly different to unity, as seen in Figure 1.

The measured pair NPI is given as

$$NPI_p(A) = \frac{\exp[-At]n_m(A) + \exp[At]n_m(-A)}{n_m(A) + n_m(-A)}. \tag{25}$$

We can now calculate the expected range for each pair NPI by linearising Eq. (25) and propagating the variance from the initial binomial distributions for each histogram bin. This is the standard error propagation method. If $\Delta NPI_p(A) = NPI_p(A) - \langle NPI_p(A)\rangle$ then to linear order

$$\Delta NPI_p(A) = \Delta n(A)\left\langle \frac{\partial NPI_p(A)}{\partial n(A)} \right\rangle + \Delta n(-A)\left\langle \frac{\partial NPI_p(A)}{\partial n(-A)} \right\rangle, \tag{26}$$

$$= \Delta n(A)\left[\frac{\exp[-At]}{\lambda(A)+\lambda(-A)} - \frac{\exp[-At]\lambda(A)}{(\lambda(A)+\lambda(-A))^2} - \frac{\exp[At]\lambda(-A)}{(\lambda(A)+\lambda(-A))^2}\right] \tag{27}$$

$$+ \Delta n(-A)\left[\frac{\exp[At]}{\lambda(A)+\lambda(-A)} - \frac{\exp[At]\lambda(-A)}{(\lambda(A)+\lambda(-A))^2} - \frac{\exp[-At]\lambda(A)}{(\lambda(A)+\lambda(-A))^2}\right]. \tag{28}$$

Using the ESFT,

$$\Delta NPI_p(A) = \Delta n(A)\left[\frac{\exp[-At]}{\lambda(A)+\lambda(-A)} - \frac{\lambda(-A)}{(\lambda(A)+\lambda(-A))^2} - \frac{\lambda(A)}{(\lambda(A)+\lambda(-A))^2}\right] \quad (29)$$

$$+ \Delta n(-A)\left[\frac{\exp[At]}{\lambda(A)+\lambda(-A)} - \frac{\lambda(A)}{(\lambda(A)+\lambda(-A))^2} - \frac{\lambda(-A)}{(\lambda(A)+\lambda(-A))^2}\right], \quad (30)$$

$$= \Delta n(A)\left[\frac{\exp[-At]-1}{\lambda(A)+\lambda(-A)}\right] + \Delta n(-A)\left[\frac{\exp[At]-1}{\lambda(A)+\lambda(-A)}\right]. \quad (31)$$

Since each histogram bin is statistically independent,

$$\left\langle \Delta NPI_p(A)^2 \right\rangle = \left\langle \Delta n(A)^2 \right\rangle \left(\frac{(\exp[-At]-1)^2}{(\lambda(A)+\lambda(-At))^2}\right) + \left\langle \Delta n(-A)^2 \right\rangle \left(\frac{(\exp[At]-1)^2}{(\lambda(A)+\lambda(-A))^2}\right). \quad (32)$$

This variance is used to calculate a range of two standard deviations around the expected value of the pair NPI, and is included in Figure 7. These limits show the increase in uncertainty as the value of A increases, and continues to increase after the error calculated using simple block averaging becomes unreliable.

We can also calculate the upper and lower limit for the pair NPI by using the extreme values for each histogram bin, with the limits that include 95% of the binomial distribution calculated in Eq. (14).These limits are included in Figure 7 as solid red lines. These limits correctly predict the increase in the noise of the pair NPI with A. The minimum possible value of the NPI of zero is also reflected by these limits. However, in the wings of the distribution the upper limit drops close to zero. This may be observed in the figure for all values of A that are larger than the value of $A = 0.85$, where the discontinuous drop of the line occurs. This sudden drop towards zero is because while the mean of the distribution describing the $-A$ histogram bin, $\lambda(-A)$, is still greater than zero, our 95% interval of the binomial distribution only includes the value zero. So, the upper and lower limits for the expected value of this histogram bin are both zero when the value $A > 0.85$. Obviously if we chose the interval to be larger than 95% this discontinuous drop would occur at a larger value of A. This mirrors what happens with the actual simulation data, where the pair NPI goes to zero when A is larger than some *critical value*. If we sampled more data the value of A where this occurs would also be larger. The exponential nature of the ESFT will mean that this increase in the *critical value* of A is extremely gradual with increased sampling.

We now turn our attention to what happens upon computing the full NPI rather than the pair NPI. To do this we have to sum up the pair NPI for each value of A as per,

$$\left\langle \exp[-\bar{\Omega}_t t] \right\rangle = \int_0^\infty dA \, \frac{n(A)+n(-A)}{N} \left\langle \exp[-\bar{\Omega}_t t] \right\rangle_A. \quad (33)$$

The pair NPI is not accurately estimated when the value of $|A|$ is too large and the value of $n(-A)$ becomes too small to get an accurate estimate. When the pair NPI with values $|A|$ that are large enough to suffer from this problem make a significant contribution to the full NPI, it will not be accurately estimated. This happens when the duration time the NPI is estimated at becomes large enough, and the larger the field F_e driving the system the earlier will be the duration at which this occurs. At these sufficiently long durations the full NPI is consistently underestimated, as seen in Figure 1.

We can measure the significance of each histogram bin in calculating the value of the NPI by calculating the value of the integrand for that bin from Eq. (4), $P_m(A) = \exp[-At]Pr(\bar{\Omega}_t = A)$. The integrand is actually the mirror of the probability distribution, as follows from Eq. (1), i.e. $P_m(A) = Pr(-A)$, hence the subscript m for mirror.

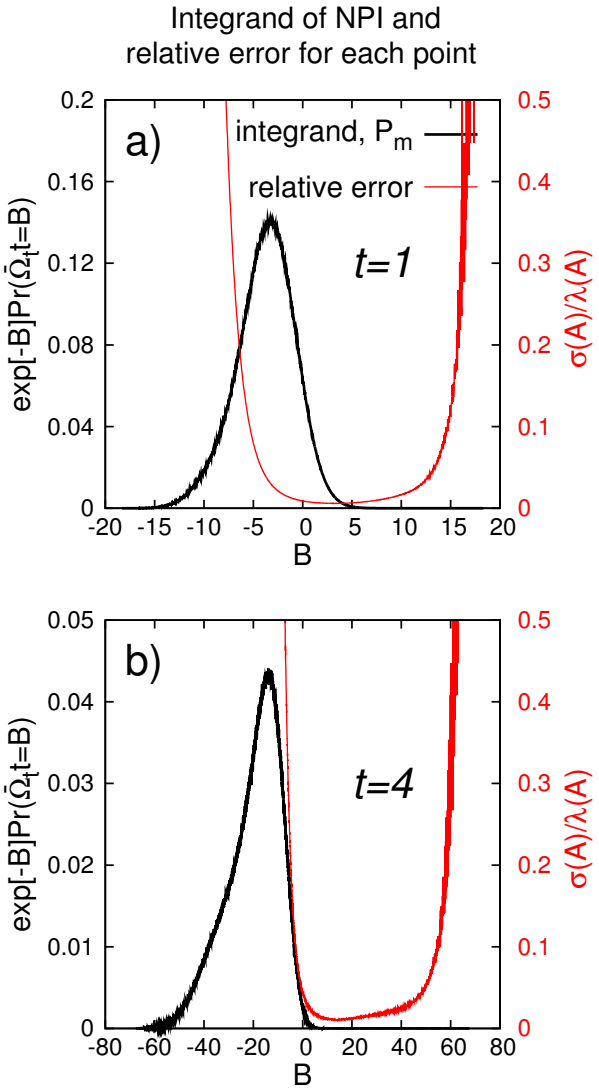

Figure 8. Integrand of the full NPI calculating shown for each histogram bin, labelled by At. Shown along side in red is the standard deviation of each bin height divided by the mean of that histogram bin, labelled as the relative error. This is plotted for the two trajectory durations, $t = 1$ on top and $t = 4$ below. Note: the dummy variable B is related to the previously used dummy variable A as per $B = At$. The histogram bin width used is 0.02.

This is plotted in Figure 8 for each histogram bin. We have also included a measure of the error in the calculated probability of each bin, the standard deviation of the bin height divided by the average height. This represents the relative error for each value of A. This data is from the same system as Figure 1, the colour conductivity system with a field strength of 2.0. The probability of each histogram bin is given by $Pr(\bar{\Omega}_t = A) = \lambda(A)/N$ where $\lambda(A)$ is determined using the ESFT in Eq. (12). The standard deviation is the square root of the variance for each histogram bin height, given by the binomial distribution, $\sigma(A) = \sqrt{\lambda(A)(1 - \lambda(A)/N)}$. Two trajectory durations are included, $t = 1$, where the calculated value for the NPI is as close to unity as expected, and $t = 4$ where the calculated value is far from unity. We can see that the most significant histogram bins in the first case have a much lower relative error than the most significant histogram bins at the latter time. This explains in detail why the NPI can be calculated in a numerical system for

short times, but fails at later times.

We have seen that negative values of the dissipation function become less likely as the trajectory length increases. We would expect the mean of $\bar{\Omega}_t t$ will scale with t, while the standard deviation will scale with $1/\sqrt{t}$. As such, at long time we can approximate the distribution of $\bar{\Omega}_t t$ with a Dirac-delta distribution. Using this approximation the value of the NPI will be

$$\int_{-\infty}^{\infty} dA\delta\left(A - \left\langle\bar{\Omega}_t\right\rangle\right)\exp[-At] = \exp\left[-\left\langle\bar{\Omega}_t\right\rangle t\right]. \tag{34}$$

This means that given enough time the measured NPI will be related to a transport coefficient, and in our case this will be the self-diffusion coefficient,

$$D = \frac{N_p - 1}{N_p}\frac{1}{\beta\rho}\lim_{t\to\infty}\lim_{F_{ex}\to 0}\frac{\langle J(t)\rangle}{F_{ex}}, \tag{35}$$

where

$$J(t) = \frac{1}{V}\sum_{i=1}^{N_p} c_i\dot{x}_i, \tag{36}$$

so

$$\langle J(t)\rangle \approx \frac{\bar{\Omega}_t}{\beta V F_{ex}}. \tag{37}$$

This approximation, Eq. (37), becomes very accurate for large enough values of t, for which the system will certainly appear to be in a steady state. Assuming a Dirac-delta distribution for $\bar{\Omega}_t t$, we can calculate $\langle J\rangle$ from the NPI as

$$\langle J(t)\rangle \approx \frac{-\ln\left\langle e^{-\bar{\Omega}_t t}\right\rangle}{\beta t V F_{ex}} \tag{38}$$

or calculate it directly as

$$\langle J(t)\rangle \approx \frac{1}{t}\left\langle\int_0^t J(s)ds\right\rangle \tag{39}$$

which will become increasingly accurate as the duration is increased. From these values of $\langle J\rangle$ we can obtain a time dependent diffusion coefficient for each case, which is plotted in Figure 9. These simulations use the same system as Figure 1 with a field strength of 2.0, but using an ensemble of 10^5 trajectories. We can see that as the trajectory duration increases, the value calculated from the NPI asymptotically approaches the direct calculation, which has become constant, demonstrating that the distribution of dissipation function values is approaching a Dirac-delta distribution. We note that significantly longer trajectories are needed to see this convergent behaviour than were necessary to see the NPI fail. Due to the large exponential functions being evaluated it was necessary to use quadruple precision in the numerical evaluation.

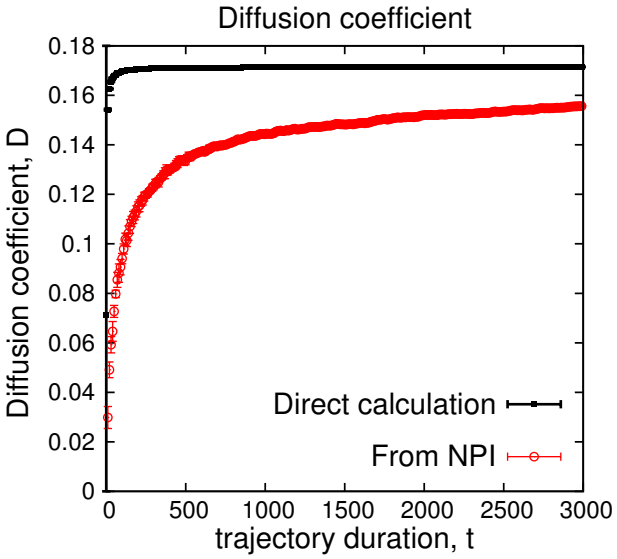

Figure 9. Long duration comparison between NPI derived diffusion coefficient, and the more directly obtained NEMD diffusion coefficient.

7. Conclusions

We have considered the average, $\left\langle \exp\left[-\bar{\Omega}_t t\right]\right\rangle$, that needs to be calculated upon implementing the NPI using computer simulations or experimental data. This average, which is taken over nonequilibrium path integrals, has some very interesting properties from a sampling perspective. Firstly we can prove that given an infinite amount of sampling it must come to a value of unity. But in this paper we have investigated what happens given a finite amount of sampling. If we take a fixed amount of sampling and extend the duration of the path integral, then the average tends towards a value that is less than unity. This is because the average is composed of rare events which are of particularly heavy weight. The longer the duration the rarer these events become, but in direct compensation their weights increase. By analysing the statistical sampling errors we have been able to demonstrate in detail how this happens and the effect it has on the average. As the time is increased still further the average asymptotes towards a value which may be determined from the relevant transport coefficient. This is due to the average value being accurately estimated upon approximating the distribution as a Dirac-delta distribution. In doing so we ignore the rare events, which are of crucial importance in obtaining the value of unity, consistent with the full ensemble average. At long enough durations the chance of observing one of these extremely rare events, with our finite amount of sampling, becomes completely improbable. Again, this improbability is exactly compensated for by the contribution the rare events make to the full ensemble average. It is this complete improbability that leads to the Dirac-delta distribution function giving the actual value that is approached using a fixed amount of sampling. On the other hand if we chose some finite duration, and then increased the number of independent trajectory segments that we sample, the value of unity would be approached. This is because upon increasing the sampling enough, we will eventually sample the rare events fully.

These findings are of great importance for the JE[12, 13], which is of the same mathematical form. Upon attempting to make use of the JE using a finite amount of sampled data, using computer simulations, or experiments, very similar behaviour will be observed. The error analysis developed in this paper could be straightforwardly adopted to the JE if one uses the irreversible work, $w_{ir}(t) = w(t) - \Delta A$, where $w(t)$ is the ordinary work used in the JE and ΔA is the change in free energy.

The JE is then given as

$$\langle \exp\left[-\beta w_{ir}(t)\right] \rangle = 1. \tag{40}$$

The fluctuation relation that corresponds to this is the CFT[4, 11]. The one key difference that needs to be accounted for with this is that unlike the ESFT, the CFT needs to consider reverse trajectory segments with reverse protocols which start from the final equilibrium. The ESFT only considers the initial protocol and distribution, and exploits a symmetry between the forward and reverse trajectory segments, that is usually not present when use is made of the CFT.

We dedicate this paper to the memory of our colleague, mentor and friend Ian Snook.

References

[1] Sevick EM, Prabhakar R, Williams SR, et al. Fluctuation Theorems. Annual review of physical chemistry. 2008;59:603–33.

[2] Evans DJ, Searles DJ. The Fluctuation Theorem. Advances in Physics. 2002 Jan;51:1529–1585.

[3] Evans DJ, Searles DJ. Equilibrium microstates which generate second law violating steady states. Physical Review E. 1994;50:1645–1648.

[4] Crooks GE. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Physical Review E. 1999;60:2721–26.

[5] Jarzynski C. Nonequilibrium work relations: foundations and applications. The European Physical Journal B. 2008;64:331–340.

[6] Seifert U. Stochastic thermodynamics, fluctuation theorems and molecular machines. Reports on Progress in Physics. 2012;75:126001.

[7] Evans DJ, Searles DJ, Williams SR. On the fluctuation theorem for the dissipation function and its connection with response theory. The Journal of chemical physics. 2008 Jan;128:014504.

[8] Evans DJ, Searles DJ, Williams SR. Erratum: "On the fluctuation theorem for the dissipation function and its connection with response theory" [J. Chem. Phys. 128, 014504 (2008)]. The Journal of Chemical Physics. 2008;128:249901.

[9] Evans DJ, Searles DJ, Williams SR. Dissipation and the relaxation to equilibrium. Journal of Statistical Mechanics: Theory and Experiment. 2009 Jul;2009:P07029.

[10] Evans DJ, Searles DJ, Williams SR. A simple mathematical proof of boltzmann's equal a priori probability hypothesis. In: Chmelik C, Kanellopoulos N, Karger J, et al., editors. Diffusion fundamentals iii. Leipziger Universitatsverlag; 2009.

[11] Crooks GE. Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. Journal of statistical physics. 1998;90:1481–1487.

[12] Jarzynski C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. Physical Review E. 1997;56:5018–5035.

[13] Jarzynski C. Nonequilibrium Equality for Free Energy Differences. Physical Review Letters. 1997; 78:2690.

[14] Carberry DM, Williams SR, Wang GM, et al. The Kawasaki identity and the Fluctuation Theorem. The Journal of chemical physics. 2004 Nov;121:8179–82.

[15] Searles DJ, Johnston BM, Evans DJ, et al. Time reversibility, correlation decay and the steady state fluctuation relation for dissipation. Entropy. 2013;15:1503–1515.

[16] Evans DJ, Morriss GP. Statistical mechanics of nonequilibrium liquids. 2nd ed. Cambridge: Cambridge University Press; 2008.

[17] Liphardt J, Dumont S, Smith SB, et al. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. Science. 2002 Jun;296:1832–5.

[18] Douarche F, Ciliberto S, Petrosyan A, et al. An experimental test of the Jarzynski equality in a mechanical experiment. Europhysics Letters. 2005;70:593–599.

[19] Bena I, Van Den Broeck C, Kawai R. Jarzynski equality for the Jepsen gas. Europhysics Letters. 2005; 71:879–885.

[20] Jayannavar AM, Sahoo M. Charged particle in a magnetic field: Jarzynski equality. Physical Review E. 2007;75:032102.
[21] West DK, Olmsted PD, Paci E. Free energy for protein folding from nonequilibrium simulations using the Jarzynski equality. Journal of Chemical Physics. 2006;125:204910.
[22] Gore J, Ritort F, Bustamante C. Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. PNAS. 2003;100:12564–12569.
[23] Reid JC, Cunning BV, Searles DJ. Different approaches for evaluating exponentially weighted nonequilibrium relations. 2010;133:154108.
[24] Chong SH, Otsuki M, Hayakawa H. Generalized Green-Kubo relation and integral fluctuation theorem for driven dissipative systems without microscopic time reversibility. Physical Review E. 2010;81:041130.
[25] Liu F, Luo YP, Huang MC, et al. A generalized integral fluctuation theorem for general jump processes. Journal of Physics A: Mathematical and Theoretical. 2009;42:332003.
[26] Speck T, Seifert U. Integral fluctuation theorem for the housekeeping heat. Journal of Physics A: Mathematical and General. 2005;38:L581–L588.
[27] Seifert U. Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem. Physical Review Letters. 2005;95:040602.
[28] Seifert U. Fluctuation theorem for birth-death or chemical master equations with time-dependent rates. Journal of Physics A: Mathematical and General. 2004;37:L517–L521.
[29] Evans DJ, Morriss GP. Nonlinear-response theory for steady planar Couette flow. Physical Review A. 1984;30:1528–1530.
[30] Evans DJ, Searles DJ. Steady states, invarient meaures, and resposne theory. Physical Review E. 1995; 52.
[31] Van Zon R, Cohen EGD. Extended heat-fluctuation theorems for a system with deterministic and stochastic forces. Physical Review E. 2004 May;69:056121.
[32] Weeks J, Chandler D, Andersen H. Role of repulsive forces in determining the equilibrium structure of simple liquids. The Journal of chemical physics. 1971;54:5237–5247.